Quantum and Fisher Information from the Husimi and Related Distributions

Paul B. Slater*
ISBER, University of California,
Santa Barbara, CA 93106

(Dated: February 1, 2008)

Abstract

The two principal/immediate influences — which we seek to interrelate here — upon the undertaking of this study are papers of Życzkowski and Słomczyński (J. Phys. A 34, 6689 [2001]) and of Petz and Sudár (J. Math. Phys. 37, 2262 [1996]). In the former work, a metric (the Monge one, specifically) over generalized Husimi distributions was employed to define a distance between two arbitrary density matrices. In the Petz-Sudár work (completing a program of Chentsov), the quantum analogue of the (classically unique) Fisher information (montone) metric of a probability simplex was extended to define an uncountable infinitude of Riemannian (also monotone) metrics on the set of positive definite density matrices. We pose here the questions of what is the specific/*unique* Fisher information metric for the (classically-defined) Husimi distributions and how does it relate to the *infinitude* of (quantum) metrics over the density matrices of Petz and Sudár? We find a highly proximate (small relative entropy) relationship between the probability distribution (the quantum Jeffreys' prior) that yields quantum universal data compression, and that which (following Clarke and Barron) gives its classical counterpart. We also investigate the Fisher information metrics corresponding to the *escort* Husimi, positive-P and certain Gaussian probability distributions, as well as, in some sense, the discrete Wigner *pseudoprobability.* The *comparative noninformativity* of prior probability distributions — recently studied by Srednicki (Phys. Rev. A 71, 052107 [2005]) — formed by normalizing the volume elements of the various information metrics, is also discussed in our context.

PACS numbers: Valid PACS 03.67.-a, 02.50.Tt, 02.50.Sk, 02.40.Ky

*Electronic address: slater@kitp.ucsb.edu

I. INTRODUCTION

The two-level quantum systems (TLQS) are describable (nonclassically) in terms of 2×2 density matrices (ρ) — Hermitian nonnegative definite matrices of trace unity. These matrices can be parametrized by points in the unit ball (Bloch ball/sphere [1, p. 10244]) in Euclidean 3-space. On the other hand, the TLQS can be described in a *classical* manner using a generalization of the Husimi distribution [2] [3, sec. 4.1] (cf. [4, 5, 6, 7, 8, 9]). "The Husimi function is a function on phase space, and takes only non-negative values while the Wigner function can be negative and is usually violently oscillating. Hence the Husimi function can be regarded as a probability distribution in phase space, and its order of delocalization can be a measure of chaoticity of quantum states" [10]. (Note that the original Husimi distribution was defined only for density operators in *separable* Hilbert space — one which admits a countable orthonormal basis — while the distribution to be studied here is defined over a finite-dimensional Hilbert space.)

There is an (uncountable) *infinitude* [11, sec. 16.7] of (quantum monotone) Riemannian metrics that can be attached to the Bloch ball of TLQS. Contrastingly, in the classical context of the Husimi distribution, there is not an infinitude, but rather a *single* distinguished (up to a constant multiple) monotone Riemannian metric — the *Fisher information* metric [12, 13, 14]. ("In the classical case, decision theory provides a unique montone metric, namely, the Fisher information. In the quantum case, there are infinitely many monotone metrics on the state space" [15, p. 2672].) So, it appears to be an question of obvious interest — which we seek to address here — of how one reconciles/deals with this phenomenon of classical uniqueness and quantum non-uniqueness, as applied to essentially the *same* objects (that is, the TLQS).

II. MONOTONE METRICS

The monotone metrics are all *stochastically monotone* [15]. That is, geodesic distances (as well as relative entropies) between density matrices *decrease* under coarse-grainings (completely positive trace-preserving maps, satisfying the Schwarz inequality: $T(a^*a) \geq T(a)^*T(a)$). These metrics can be used for purposes of statistical distinguishability [15]. The monotone metrics for the TLQS have been found to be rotationally invariant

over the Bloch ball, depending only on the radial coordinate r, that is the distance of the state in question from the origin $(0,0,0)$ — corresponding to the fully mixed state. They are splittable into radial and tangential components of the form [15, eq. (3.17)],

$$ds^2_{monotone} = \frac{1}{1-r^2}dr^2 + \Big((1+r)f(\frac{1-r}{1+r})\Big)^{-1} dn^2. \tag{1}$$

Here, using spherical coordinates (r, θ_1, θ_2), one has $dn^2 = r^2 d\theta_1^2 + r^2 \sin^2\theta_1 d\theta_2^2$. Further, $f : \mathbb{R}^+ \to \mathbb{R}^+$ is an operator monotone function such that $f(t) = tf(t^{-1})$ for every $t > 0$. (A function is operator monotone if the relation $0 \le K \le H$, meaning that $H - K$ is nonnegative definite, implies $0 \le f(K) \le f(H)$ for any such matrices K and H of any order.) The radial component is *independent* of the function f, and in the case of the Bures (minimal monotone) metric (corresponding to the particular choice $f_{Bures}(t) = \frac{1+t}{2}$), the tangential component is independent of r [16].

In the classical context of the Husimi distribution, there is not an infinitude, but rather a *single* distinguished (up to a constant multiple) monotone metric — the *Fisher information* metric [12, 13, 14]. (The counterpart here to stochastic mappings — which are the appropriate morphisms in the category of quantum state spaces — are stochastic *matrices* [15].) The ij-entry of the Fisher information matrix (tensor) is the expected value with respect to the probability distribution in question of the product of the *first* derivative of the logarithm of the probability with respect to its i-th parameter times the analogous first derivative with respect to its j-th parameter. (Under certain regularity conditions, the Fisher information matrix is equal to the "second derivative matrix for the informational divergence (relative entropy)" [17, pp. 455-456], [18, p. 43].) The volume element of the Fisher information metric can be considered — in the framework of Bayesian theory — as a prior distribution (Jeffreys' prior [17, 19, 20]) over, for our purposes here, the Bloch ball of TLQS.

A. Fisher information metric for the Husimi distribution

We have found (having to make use of numerical, as well as symbolic MATHEMATICA procedures in our quest) that for the Husimi distribution over the TLQS, the Fisher information metric takes the specific form (cf. (2)),

$$ds^2_{Fisher_{Hus}} = \frac{-2r - \log(\frac{1-r}{1+r})}{2r^3} dr^2 + \Big((1+r)f_{Hus}(\frac{1-r}{1+r})\Big)^{-1} dn^2. \tag{2}$$

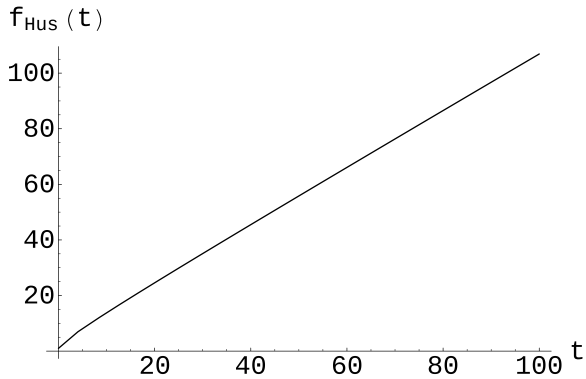

FIG. 1: The monotone function $f_{Hus}(t)$ that yields the *tangential* component of the Fisher information metric over the trivariate Husimi probability distributions for the two-level quantum systems.

Here,

$$f_{Hus}(t) = \frac{(t-1)^3}{t^2 - 2t\log t - 1}. \tag{3}$$

Now, a plot (Fig. 1) shows $f_{Hus}(t)$ to be, in fact, a *monotone* function. ($f_{Hus}(t)$ is "almost" equal to $\frac{(t-1)^3}{t^2-2t-1} = t-1$.) It has a singularity at $t=1$, corresponding to the fully mixed state ($r=0$), where $f_{Hus}(1+\Delta t) \approx 3 + 3\Delta t/2$, though we have not attempted to confirm its *operator* monotonicity. Also, $f_{Hus}(t)$ fulfills the self-adjointness condition $f(t) = tf(t^{-1})$ of Petz and Sudár [15, p. 2667], at least at $t \neq 1$. For the pure states, that is $t = 0, r = 1$, we have $\lim_{t\to 0} f_{Hus}(t) = 1$.

We further have the relation,

$$c_{Hus}(p,q) = \frac{1}{qf_{Hus}(\frac{p}{q})} = \frac{q^2 - p^2 - 2pq\log\frac{q}{p}}{(q-p)^3}, \tag{4}$$

where $c_{Hus}(p,q)$ is a specific "Morozova-Chentsov" function. There exist one-to-one correspondences between Morozova-Chentsov functions, monotone metrics and operator means [21, Cor. 6]. "Operator means are binary operations on positive operators which fulfill the main requirements of monotonicity and the transformer inequality" [21].

We can write (1) more explicitly as

$$ds^2_{Fisher_{Hus}} = \frac{-2r - \log(\frac{1-r}{1+r})}{2r^3}dr^2 + \frac{2r \ + (1-r^2)\log(\frac{1-r}{1+r})}{4r^3}dn^2. \tag{5}$$

Certainly, $ds^2_{Fisher_{Hus}}$ does not have — in terms of the radial component — the specific form (1) required of a monotone metric (cf. [22]). In Fig. 2 we show both the *radial*

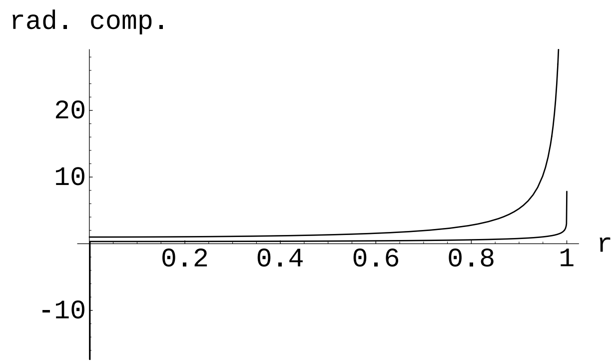

FIG. 2: The radial components of *any* monotone metric and that of the Fisher information metric derived from the family of trivariate Husimi distributions over the TLQS. The one for the (nondenumerably infinite) class $ds^2_{monotone}$ dominates that for $ds^2_{Fisher_{Hus}}$.

components of (any) $ds^2_{monotone}$ and of $ds^2_{Fisher_{Hus}}$. Petz [23, p. 934] attributes the unvarying nature ($\frac{1}{1-r^2}$) of the radial component of the (quantum) monotone metrics to the (classical) Chentsov uniqueness (of Fisher information) theorem [12, 13]. "Loosely speaking, the unicity [sic] result in the [probability] simplex case survives along the diagonal and the off-diagonal provides new possibilities for the definition of a stochastically invariant metric" [15, p. 2664].

If we (counterfactually) equate the volume element of $ds^2_{Fisher_{Hus}}$ to that of a generic monotone metric (1), and solve for $f(t)$, we obtain a monotonically-*decreasing* function (Fig. 3) (cf. [22]),

$$f_{counterfactual}(t) = \frac{\sqrt{2}\,(-1+t)^{\frac{9}{2}}}{t\ (1+t)\ \sqrt{\left(-1+t^2-2\,t\ \log(t)\right)^2\ \left(2-2\,t+(1+t)\ \log(t)\right)}}. \tag{6}$$

Converting to cartesian coordinates (x, y, z), the *trace* of $ds^2_{Fisher_{Hus}}$ can be simply expressed as $-\log\left(\frac{1-R}{1+R}\right)/(2R)$, where $R = \sqrt{x^2+y^2+z^2}$ (cf. [14, 24]). Also, at the fully mixed state $(x = y = z = 0)$, the metric is simply *flat*, that is

$$ds^2_{Fisher_{Hus}} = \frac{1}{3}(dx^2 + dy^2 + dz^2). \tag{7}$$

(The Riemann and Ricci tensors evaluated at the fully mixed state have no non-zero entries.)

Numerical evidence indicates that the Fisher information matrix for the Husimi distribution over the TLQS is bounded by the corresponding information matrices for the (quantum) monotone metrics, in the sense that the monotone metric tensors minus the Fisher-Husimi information tensor are positive definite.

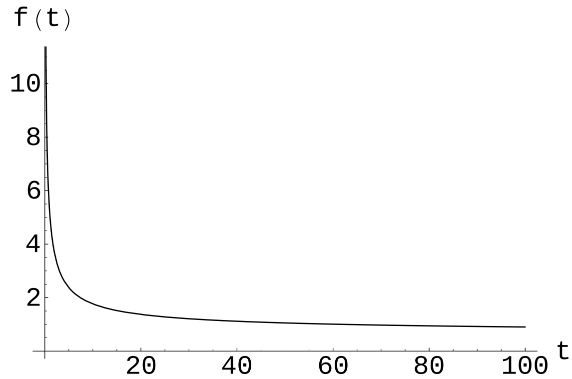

FIG. 3: Monotonically-*decreasing* function $f_{counterfactual}$ obtained by equating the volume element of $ds^2_{Fisher_{Hus}}$ to that of a generic monotone metric (1)

We can normalize the volume element of $ds^2_{Fisher_{Hus}}$ to a probability distribution p_{Hus} by dividing by the Fisher information metric *volume* ≈ 1.39350989367660. If we generate a "hybridized-Husimi" (quantum [15]) monotone metric, $ds^2_{HYB_{Hus}}$, *via* the formula (1), using $f_{Hus}(t)$, then the volume of the Bloch ball of TLQS in terms of this newly-generated monotone metric is $\frac{1}{2}\pi^2(4-\pi) \approx 4.23607 > 1.39351$. Using this as a normalization factor, we obtain a probability distribution ($p_{HYB_{Hus}}$) of interest over the TLQS.

III. COMPARATIVE NONINFORMATIVITIES

Let us compare p_{Hus} — in the manner employed in [25, 26] (cf. [27, 28, sec. VI]) — with the prior probability distribution (p_{Bures}). The latter is gotten by normalizing the volume element of the well-studied *minimal* monotone (Bures) metric ([29, eq. (7)] [30, eq. (16)]), that is,

$$p_{Bures} = \frac{r^2 \sin\theta_1}{\pi^2\sqrt{1-r^2}}, \tag{8}$$

generated from (1) using the operator monotone function $f_{Bures}(t) = \frac{1+t}{2}$. (We avoid the specific designations $f_{min}(t)$ and $f_{max}(t)$ because these are usually, confusingly, considered to generate the maximal and minimal monotone metrics, respectively [15, eq. (3.21)]. Our integrations of probability distributions are conducted over $r \in [0,1], \theta_1 \in [0,\pi]$ and $\theta_2 \in [0, 2\pi]$.)

The *relative entropy* (Kullback-Leibler distance) of p_{Bures} with respect to p_{Hus} [which we denote $S_{KL}(p_{Bures}, p_{Hus})$] — that is, the *expected* value with respect to p_{Bures} of $\log \frac{p_{Bures}}{p_{Hus}}$ —

is 0.130845 "nats" of information. (We use the *natural* logarithm, and not 2 as a base, with one nat equalling 0.531 bits.) Let us note that the Shannon entropy ($S_{Shannon}$) of the Husimi distribution is the Wehrl entropy (S_{Wehrl}) of the corresponding quantum state. Explicitly implementing [31, eq. (6)], we have for the TLQS,

$$S_{Wehrl} = \frac{1}{4r}\Big(2r + 4r\log 2 + (1+r^2)\log\big(\frac{1-r}{1+r}\big) - 2r\log\left(1-r^2\right)\Big). \tag{9}$$

S_{Wehrl} is always greater than the von Neumann entropy, $S_{vN} = -\mathrm{Tr}\rho\ln\rho$, which for the TLQS is expressible as

$$S_{vN} = \frac{1}{2}\Big(2\log 2 + r\log\big(\frac{1-r}{1+r}\big) - \log\left(1-r^2\right)\Big). \tag{10}$$

(We, of course, notice the omnipresence in these last two formulas, as well as in (5) and further formulas below of the term $W \equiv \log\left(\frac{1-r}{1+r}\right)$. The two eigenvalues $[\lambda_1, \lambda_2 = 1-\lambda_1]$ of ρ are $\frac{1\pm r}{2}$, so W is expressible as $\log\left(\frac{\lambda_1}{\lambda_2}\right)$.) Each monotone metric can be obtained in the form of a "contrast functional" for a certain convex subset of relative entropies [32, 33].

A. Bures prior

Now, let us convert p_{Bures} to a *posterior* probability distribution ($post_{Bures}$) by assuming the performance of *six* measurements, *two* (with one outcome "up" and the other "down") in each of the $x-$, y- and z-directions. Normalizing the product of the prior p_{Bures} and the *likelihood* function corresponding to the six measurement outcomes [25, p. 3],

$$post_{Bures} = \frac{192 p_{Bures}(1-x^2)(1-y^2)(1-z^2)}{71}, \tag{11}$$

we find $S_{KL}(post_{Bures}, p_{Hus}) = 0.0912313 < 0.130845$. (The cartesian coordinates in (11) are transformed to the spherical ones employed in our analysis.) So, in this sense p_{Bures} is *more* noninformative than p_{Hus}, the relative entropy being *reduced* by *adding* information to p_{Bures}. On the other hand, p_{Bures} — corresponding to the *minimal* monotone metric — is itself the *least* noninformative of the monotone-metric priors ($p_{monotone}$) [25]. (Luo has established an inequality between the [monotone metric] Wigner-Yanase *skew information* and its minimal monotone counterpart [34].)

Reversing the arguments of the relative entropy functional, we obtain $S_{KL}(p_{Hus}, p_{Bures}) =$ .0818197. But now, following the same form of posterior construction, we find

$S_{KL}(post_{Hus}, p_{Bures}) = 0.290405 > .0818197$, further supportive of the conclusion that p_{Bures} is *more* noninformative than p_{Hus}. In some sense, then, p_{Bures} assumes *less* about the data than p_{Hus}. But this diminishability of the relative entropy is limited. If we convert p_{Bures} to a new posterior $Post_{Bures}$ using the *square* of the likelihood function above — that is, assuming *twelve* measurements, *four* (with two outcomes "up" and the other two "down") in each of the $x-$, y- and z-directions, giving

$$Post_{Bures} = \frac{21504 p_{Bures}[(1-x^2)(1-y^2)(1-z^2)]^m}{3793}, m = 2, \tag{12}$$

then $S_{KL}(Post_{Bures}, p_{Hus}) = 0.292596 > 0.130845$. To much the same effect, if we use a likelihood based on the *optimal/nonseparable* set of measurements for *two* qubits, consisting of five possible measurement outcomes, given in [35, eq. (8)], to convert p_{Bures} to a new posterior, then the relative entropy reaches higher still, that is from 0.130845 to 0.623855. (Employing a likelihood based on the optimal/nonseparable set of measurements for *three* qubits, consisting of eight possible measurement outcomes [35, eq. (9)], the relative entropy with respect to p_{Hus} increases further to 1.51365.) Actually, if we *formally* take $m = \frac{1}{2}$ in eq. (12), and renormalize to a new posterior, we obtain a superior reduction, that is, to $0.07167 < 0.0912313$. (Further, with $m = \frac{5}{8}$, we get 0.0702389 and 0.0732039, with $m = \frac{3}{4}$.)

B. Morozova-Chentsov prior

In [25], it was found that the ("Morozova-Chentsov") prior distribution,

$$p_{MC} = \frac{.00513299[\log\left(\frac{1-r}{1+r}\right)]^2 \sin\theta_1}{\sqrt{1-r^2}}, \tag{13}$$

that is, the normalized volume element of the monotone metric (1) based on the operator monotone function,

$$f_{MC}(t) = \frac{2(t-1)^2}{(1+t)(\log t)^2}, \tag{14}$$

was apparently the *most* noninformative of those (normalizable) priors based on the operator monotone functions that had been explicitly discussed in the literature. Now, $S_{KL}(p_{MC}, p_{Hus}) = 1.37991$, that is, quite large. This can be reduced to 0.893996 if, into p_{MC}, one incorporates $m = 6$ measurements of the type described above; diminished further to 0.561901 with $m = 12$; and further still to 0.471852 — the greatest reduction of this type — with $m = 18$. (For $m = 24$, it starts to rise to 0.652441.)

But, if we again use the likelihood based on the optimal nonseparable measurement of two qubits [25, eq. (8)], with just five measurements, the relative entropy of the corresponding posterior form of p_{MC} with respect to p_{Hus} is reduced to 0.342124, which is the *smallest* we have achieved so far along these lines. (For the mentioned optimal nonseparable measurement scheme for *three* qubits, the reduction is quite minor, only to 1.33492 nats.) We obtained intermediate-sized reductions to 0.45524 and 0.492979, respectively, by using for our measurements, *twenty* projectors oriented to the vertices [36, secs. 9, 10] of a dodecahedron and of an icosahedron. (The primary measurement scheme used above, and in [25], with six measurements oriented along three orthogonal directions, is tantamount to the use of an *octahedron*.)

C. Hilbert-Schmidt prior

The prior distribution generated by normalizing the volume element of the Hilbert-Schmidt metric over the Bloch sphere is [25, eq. (10)] [16, eq. (31)]

$$p_{HS} = 3\frac{r^2 \sin\theta_1}{4\pi}, \tag{15}$$

which is simply the uniform distribution over the unit ball. The Hilbert-Schmidt volume element can be reproduced using the formula (1) for a quantum monotone metric, making use of $f_{HS} = \frac{(1+t)^2}{\sqrt{t}}$, but this function is neither monotone-increasing nor decreasing over $t \in [0,1]$ (cf. [37]).

We have that $S_{KL}(p_{Hus}, p_{HS}) = 0.0579239$ and $S_{KL}(p_{HS}, p_{Hus}) = 0.05443$. Now, in terms of our usual posterior distributions based on six measurements, $S_{KL}(post_{Hus}, p_{HS}) = 0.0236596$ and $S_{KL}(post_{HS}, p_{Hus}) = 0.278953$, so we can conclude that the Husimi prior p_{Hus} is more noninformative than the Hilbert-Schmidt prior p_{HS}.

IV. UNIVERSAL DATA COMPRESSION

Employing p_{Hus} as a prior distribution (Jeffreys' prior) over the family (Riemannian manifold) of Husimi qubit probability distributions, the (classical) *asymptotic minimax/maximin redundancy* of *universal data compression* is equal to [18, eq. (2.4)] [17],

$$\frac{3}{2}\log\frac{n}{2\pi e} + \log 1.39350989367660 = \frac{3}{2}\log\frac{n}{2\pi e} + 0.331826 = \frac{3}{2}\log n - 3.92499, \tag{16}$$

where n is the sample size (the number of qubits [TLQS]) and we used the before-mentioned volume of $ds^2_{Fisher_{Hus}}$. ("Suppose that X is a discrete random variable whose distribution is in the parametric family $\{P_\theta : \theta \in \Theta\}$ and we want to encode a block of data for transmission. It is known that a lower bound on the expected codeword length is the entropy of the distribution. Moreover, this entropy bound can be achieved, within one bit, when the distribution is known. Universal codes have expected length near the entropy no matter which member of the parametric family is true. The redundancy of a code is defined to be the difference between its expected length and its entropy" [17, p. 459].)

For the *quantum*/nonclassical counterpart [38] (cf. [39, 40, 41]), let us consider the use of the "Grosse-Krattenthaler-Slater" ("quasi-Bures") probability distribution [35, eq. (33)],

$$p_{GKS} = \frac{0.0832258e}{1-r^2}\Big(\frac{1-r}{1+r}\Big)^{\frac{1}{2r}} r^2 \sin\theta_1. \tag{17}$$

This is the normalized form of the monotone metric (1) associated with the (presumably operator) monotone function,

$$f_{GKS}(t) = \frac{t^{t/(t-1)}}{e}. \tag{18}$$

(Taking limits, we have for the fully mixed state, $f_{GKS}(1) = 1$ and for the pure states, $f_{GKS}(0) = e^{-1}$.) It appears [42] (though not yet fully rigorously established) that the (quantum) asymptotic minimax/maximin redundancy, employing p_{GKS} as a prior probability distribution over the 2×2 density matrices (*and* their n-fold tensor products (cf. [43])), is $\frac{3}{2}\log n - 1.77062$. This is *greater* than the classical (Husimi-Fisher-information-based) analog (16) by 2.20095 nats of information. It would seem that this difference is attributable to the greater dimensionality (2^n) of an n-qubit Hilbert space, as opposed to a dimensionality of $3n$ for n trivariate Husimi probability distributions over the TLQS.

We further note that $S_{KL}(p_{Bures}, p_{HYB_{Hus}}) = 0.00636046$ and $S_{KL}(p_{HYB_{Hus}}, p_{Bures}) = 0.0062714$, both being very small. Smaller still, $S_{KL}(p_{Bures}, p_{GKS}) = 0.00359093$ and $S_{KL}(p_{GKS}, p_{Bures}) = 0.00354579$ — whence the designation $p_{quasi-Bures} \equiv p_{GKS}$. But then, even more strikingly, we computed that $S_{KL}(p_{GKS}, p_{HYB_{Hus}}) = 0.000397852$ and $S_{KL}(p_{HYB_{Hus}}, p_{GKS}) = 0.000396915$. In Fig. 4 we show the *one*-dimensional marginal probability distributions over the radial coordinate r of the five distributions p_{Bures}, $p_{HYB_{Hus}}$, p_{Hus}, p_{GKS} and p_{MC}, with those for $p_{HYB_{Hus}}$ and p_{GKS} being — as indicated — particularly proximate.

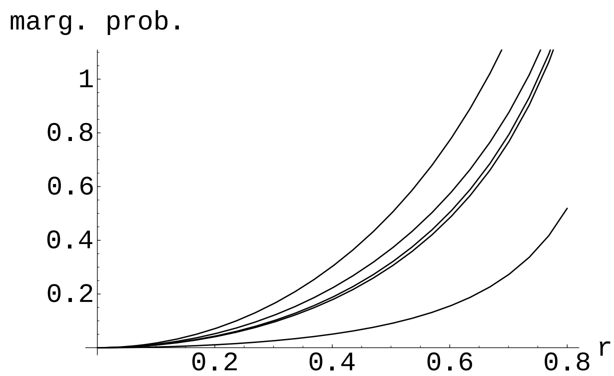

FIG. 4: Plots of one-dimensional marginal probability distributions over the radial coordinate r of p_{Bures}, $p_{HYB_{Hus}}$, p_{GKS}, p_{Hus} and p_{MC}. The order of dominance of the curves is: $p_{Hus} > p_{Bures} > p_{GKS} > p_{HYB_{Hus}} > p_{MC}$. The marginal distributions of $p_{HYB_{Hus}}$ and p_{GKS} are quite close, as reflected in their small relative entropy ($\approx .0004$).

Substitution of $p_{HYB_{Hus}}$ for p_{GKS} into the quantum asymptotic (maximin) redundancy formula that has to be *maximized* over all possible prior probability distributions [42, eq. (4.3)],

$$\frac{3}{2}\log n - \frac{1}{2} - \frac{3}{2}\log 2 - \frac{3}{2}\log\pi + 4\pi\int_0^1 \left(-\log(1-r^2) + \frac{1}{2r}\log\left(\frac{1-r}{1+r}\right) - \log w(r)\right) r^2 w(r)\, dr, \quad (19)$$

leads to a very slightly decreased (and hence suboptimal) redundancy, $\frac{3}{2}\log n - 1.77101$ *vs.* $\frac{3}{2}\log n - 1.77062$. (Use of p_{Bures} as a quantum prior over the 2×2 density matrices gives us a constant term of -1.77421, use of p_{Hus}, -1.88279 and use of p_{MC}, -2.15667.) To obtain the appropriate form of $w(r)$ to use in (19), we take our probability distributions (such as (8) and (13)), divide them by $4\pi r^2$ and integrate the results over $\theta_1 \in [0,\pi]$ and $\theta_2 \in [0, 2\pi]$. (Thus, we must have $4\pi\int_0^1 w(r) r^2 dr = 1$.) The *minimax* objective function is

$$\min_{w} \max_{0\le r\le 1} \left(\frac{3}{2}\log n - \frac{1}{2} - \frac{3}{2}\log 2 - \frac{3}{2}\log\pi - \log(1-r^2) + \frac{1}{2r}\log\left(\frac{1-r}{1+r}\right) - \log w(r)\right). \quad (20)$$

The minimax is also achieved using the $w(r)$ formed from p_{GKS}.

We can, additionally, achieve an extremely good fit to p_{Hus} by proceeding in somewhat an *opposite* fashion to that above — *reversing* our hybridization procedure. Employing $f_{GKS}(t)$, rather than $f_{Hus}(t)$ in the expression (2) for $ds^2_{Fisher_{Hus}}$ and obtaining the corresponding normalized (dividing by 4.00277) volume element ($p_{H\tilde{Y}B_{GKS}}$), we find

$S_{KL}(p_{H\tilde{Y}B_{GKS}}, p_{Hus}) = 0.000316927$. (Interchanging the arguments of the relative entropy functional, we get 0.000317754.) It is quite surprising, then, that a joint plot of $f_{GKS}(t)$ and $f_{Hus}(t)$ readily shows them to be substantially *different* in character (for example, $f_{Hus}(50) = 55.8161$ and $f_{GKS}(50) = 19.9227$), since they have been shown here to generate two pairs of such highly similar probability distributions, one pair composed of (quantum) monotone (p_{GKS} and $p_{HYB_{Hus}}$), and the other pair of (quantum) non-monotone metrics ($p_{H\tilde{Y}B_{GKS}}$ and p_{Hus}).

V. ESCORT-HUSIMI DISTRIBUTIONS

For the *escort*-Husimi distributions [44], we raise the probability element of the Husimi distribution to the q-th power, and renormalize to a new probability distribution. (Of course, the Husimi distribution itself corresponds to $q = 1$. If we set $\alpha = 2q - 1$, we recover the α-family of Amari [33, 45, 46].) To normalize the q-th power of the Husimi distribution, one must divide by

$$\frac{2^{-q}\Big(-(1-r)^{1+q} + (1+r)^{1+q}\Big)}{r + qr}. \tag{21}$$

A. The case $q = 2$

For (entropic index) $q = 2$, the Fisher information metric takes the form

$$ds^2_{Fisher_{q=2}} = \frac{12}{(3+r^2)^2} dr^2 + \Big((1+r) f_{q=2}(\frac{1-r}{1+r})\Big) dn^2, \tag{22}$$

where

$$f_{q=2}(t) = \frac{t^2 + t + 1}{2(t+1)}. \tag{23}$$

We have $f_{q=2}(1) = \frac{3}{4}$ and $f_{q=2}(0) = \frac{1}{2}$.

1. *Relative entropies*

Further, the relative entropies $S_{KL}(p_{Hus}, p_{Esc_{q=2}}) = 0.0114308$ and $S_{KL}(p_{Bures}, p_{Esc_{q=2}}) = 0.42964$, So, it appears that $p_{Esc_{q=2}}$ is even less noninformative than p_{Hus} (recalling that $S_{KL}(p_{Bures}, p_{Hus}) = 0.130845 < 0.42964$), which in turn we found above was less noninformative than the prior probabilities formed from any of the (quantum) monotone metrics. We

also note that $S_{KL}(post_{Bures}, p_{Esc_{q=2}}) = 0.125159 < 0.42964$. If we "hybridize" $ds^2_{Fisher_{q=2}}$ by modifying its radial component into that required of a (quantum) monotone metric, then we find that $S_{KL}(p_{Bures}, p_{HYB_{q=2}}) = 0.00246031(< (S_{KL}(p_{Bures}, p_{HYB_{Hus}}) = 0.00636046)$ is quite small.

B. The cases $q > 2$

For the escort-Husimi probability distribution with $q = 3$, the Fisher information metric takes the form

$$ds^2_{Fisher_{q=3}} = \frac{3-r^2}{(1+r^2)^2}dr^2 + \Big((1+r)f_{q=3}(\frac{1-r}{1+r})\Big)dn^2, \tag{24}$$

where

$$f_{q=3}(t) = \frac{t^2+1}{3(t+1)}. \tag{25}$$

Now, $f_{q=3}(1) = f_{q=3}(0) = \frac{1}{3}$ and a plot of $f_{q=3}(t)$ clearly manifests monotonic behavior also. (The monotonically-decreasing scalar curvature of $ds^2_{Fisher_{q=3}}$ equals $\frac{4}{3}$ at $r = 0$.) We have that $S_{KL}(p_{Bures}, p_{Esc_{q=3}}) = 0.63705 > S_{KL}(p_{Bures}, p_{Esc_{q=2}}) = 0.42964$, so the informativity (noninformativity) of the escort-Husimi prior probabilities *appears* to increase (decrease) with q.

For $q = 4$,

$$ds^2_{Fisher_{q=4}} = \frac{80(5-2r^2+r^4)}{3(5+10r^2+r^4)^2}dr^2 + \Big((1+r)f_{q=4}(\frac{1-r}{1+r})\Big)^{-1}dn^2, \tag{26}$$

where

$$f_{q=4}(t) = \frac{3(t^4+t^3+t^2+t+1)}{4(t+1)(3t^2+4t+3)}. \tag{27}$$

For $q = 5$,

$$ds^2_{Fisher_{q=5}} = \frac{3(5-r^2)(5+3r^4)}{(3+10r^2+3r^4)^2}dr^2 + \Big((1+r)f_{q=5}(\frac{1-r}{1+r})\Big)^{-1}dn^2, \tag{28}$$

where

$$f_{q=5}(t) = \frac{2(t^4+t^2+1)}{5(t+1)(2t^2+t+2)}. \tag{29}$$

We have (as found by C. Krattenthaler, making use of explicit MATHEMATICA computations of ours for $q = 2, 3, \cdots, 40$) (cf. [47, sec. 3.2] [48]),

$$f_q(t) = \frac{(q-1)\Sigma_{i=0}^{q} t^i}{q(t+1)\Sigma_{i=1}^{q-1} i(q-i)t^{i-1}}. \tag{30}$$

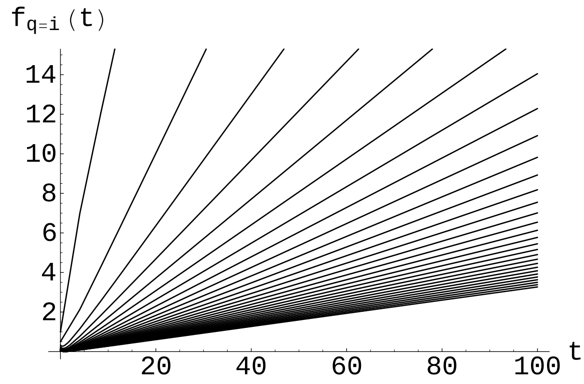

FIG. 5: The monotone functions $f_{q=i}(t), i = 1, \cdots, 30$ that yield the *tangential* components of the Fisher information metric over the escort-Husimi ($q = i$) probability distributions. The steepness of the graphs decreases as q increases

(For *odd* q some simplification in the resulting expression occurs due to cancellation by a factor of $(t + 1)$.)

In Fig. 5 we plot $f_{q=i}(t), i = 1, \cdots, 30$, revealing their common monotonically-increasing behavior. (Of course, we have $f_{q=1}(t) \equiv f_{Hus}(t)$, shown already in Fig. 1. The steepness of the curves *decreases* with increasing q.)

Let us further note that in addition to $S_{KL}(p_{Bures}, p_{HYB_{Hus}}) = 0.00636046$ and $S_{KL}(p_{Bures}, p_{HYB_{q=2}}) = 0.00246043$, we have $S_{KL}(p_{Bures}, p_{HYB_{q=3}}) = 0.0132258$, $S_{KL}(p_{Bures}, p_{HYB_{q=4}}) = 0.0238858$ and $S_{KL}(p_{Bures}, p_{HYB_{q=5}}) = 0.0327578$. (We have also been able to compute that $S_{KL}(p_{Bures}, p_{HYB_{q=1000}}) = 0.0969315$ and $S_{KL}(p_{GKS}, p_{HYB_{q=1000}}) = 0.127027$.) So, the best of these fits of p_{Bures} to the prior probabilities for the hybridized-escort-Husimi probability distributions is for $q = 2$.

C. Tangential components

Now, we can reexpress the formula (30) *without* summations, making use of the binomial theorem, as

$$f_q(t) = \frac{(-1+q)\ (-1+t)^2\ (-1+t^{1+q})}{q\ (1+t)\ (1-q+t+q\,t-t^q-q\,t^q-t^{1+q}+q\,t^{1+q})}. \tag{31}$$

So, we could study hybridized escort-Husimi metrics based on *non*-integral q using this formula. (We note that (31), in fact, yields $\lim_{q\to 1} f_q(t) \equiv f_{Hus}(t)$.) For example,

$$f_{q=\frac{1}{2}}(t) = 6 + 6\sqrt{t} + 2t - \frac{4}{1+t}. \tag{32}$$

Thus, (31) gives us (following the formulation (1)) the tangential components of the escort-Husimi Fisher information metrics for arbitrary q. (Pennini and Plastino [44] have argued, though, that in a *quantal* regime, q can be no less than 1. Tsallis statistics with an entropic index of $q = \frac{3}{2}$, Beck has contended, correctly describes the small-scale statistics of Lagrangian turbulence [49].)

D. Radial components

We do not have, at this point, a comparable complete formula for the *radial* components. However, C. Krattenthaler has shown — making use of explicit computations of ours for the cases $q = 2, 3, \cdots, 18$ — that the *denominators* of the functions giving the radial components are simply proportional to

$$u(q) = \left(\Sigma_{i=0}^{q}[\frac{Pochhammer[q-2i+1, 2i+1]r^{2i}}{2(2i+1)!}\right)^2. \tag{33}$$

(The Pochhammer symbol is synonymous with the rising or ascending factorial. The obtaining of comparable formulas for the *numerators* of the radial components might be possible using the "Rate.m" program available from the website of Krattenthaler [http://www.mat.univie.ac.at∼kratt/], if we had available additional explicit computations beyond the $q = 18$.) As way of illustration, the radial component of $ds^2_{Fisher_{q=8}}$ is expressible as

$$\frac{144\ (21 + 42\,r^2 + 135\,r^4 + 28\,r^6 + 35\,r^8 - 6\,r^{10} + r^{12})}{7u(8)}. \tag{34}$$

VI. POSITIVE P-REPRESENTATION FOR TLQS

Braunstein, Caves and Milburn focused on a specific choice of *positive* P-representation which they called the canonical form and which is always well defined [50, eq. (3.3)] (cf. [51, sec. 6.4]):

$$P_{can}(\alpha, \beta^*) \equiv \frac{1}{4\pi^2} \exp\left(-\frac{1}{4}|\alpha-\beta|^2\right) < \frac{1}{2}(\alpha+\beta)|\hat{\rho}|\frac{1}{2}(\alpha+\beta) >= \frac{1}{4\pi^2} \exp\left(-\frac{1}{4}|\alpha-\beta|^2\right) Q\left(\frac{1}{2}(\alpha+\beta)\right). \tag{35}$$

"The canonical form is clearly positive, and...it is essentially the Q-function [Husimi distribution]" [50].

We sought to implement this model, choosing for α and β *independent* 2-dimensional representations of the spin-$\frac{1}{2}$ coherent states (while for the Husimi distribution or Q-function, only, say α, need be employed). (The "positive P representation achieves [its] considerable success by doubling the number of degrees of freedom of the system, i. e., doubling the number of dimensions of the phase space" [50, p. 1153]. More typically, in the positive P-representation, α and β are allowed to vary independently over the *entire* complex plane.) However, then our result — using this choice of α and β — was *not* normalized to a probability distribution in the manner indicated in (35).

We noted that Braunstein, Caves and Milburn had commented that a "positive P representation can be defined for a large class of operators. We restrict ourselves here to those that are built up from the standard annihilation and creation operators of a harmonic oscillator. In particular, our work does not apply to generalizations of the positive P representation that include spin or pseudospin operators often used to describe a two-level atom" [50, p. 1155]. (We are not aware, however, of any specific applications reported in the literature of the positive P-representation to n-level [finite-dimensional] quantum systems.)

We did not perceive how to exactly (re)normalize the distribution (35) for our particular choices of α and β. So, we expanded just the exponential term of (35) into a power series in third order in the four *phase* variables and exactly normalized the product of this series with the remaining unmodified factor (the Q-function or Husimi distribution) to obtain a new (presumed) probability distribution. We then fit (numerically) the resultant tangential component of the associated Fisher information metric to the form (1) required of a monotone metric. In Fig. 6 we show what we (gratifyingly) obtained in this manner for $f_P(t)$. In Fig. 7 we show an approximation to the radial component of $ds^2_{Fisher_P}$, similarly obtained. (The positive P-function "seems to possess some interesting properties and may deserve close inspection"[52, p. 175].) It would be of interest to see how near the associated probability distributions (p_P and p_{HYB_P}) would be to the probability distributions (already discussed above) $p_{GKS}, p_{Hus}, p_{HYB_{Hus}}$ and $p_{H\tilde{Y}B_{GKS}}$. Most pressing, though, is the question of whether or not the concept of a positive P-representation does, in fact, have a meaningful and natural theoretical application to the n-level quantum systems.

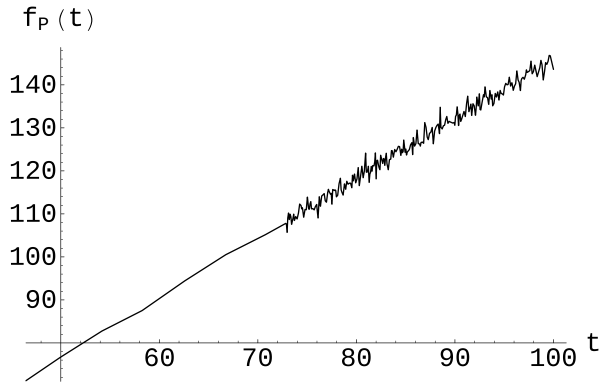

FIG. 6: *Approximation* to the presumed operator monotone function $f_P(t)$ yielding the *tangential* component of $ds^2_{Fisher_P}$ for the positive P-representation over the two-level quantum systems

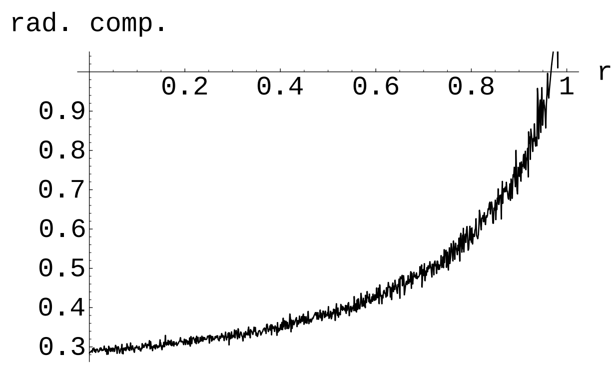

FIG. 7: *Approximation* to the *radial* component of $ds^2_{Fisher_P}$ for the positive P-representation over the two-level quantum systems

VII. GAUSSIAN DISTRIBUTION

An approach quite distinct from that of the Husimi probability distributions, but still *classical* in nature, to modeling quantum systems has been presented in [4, 5, 6, 7, 8] (cf. [9]). Here the family of probability distributions is taken as that of the Gaussian (complex multivariate normal distributions) having covariance matrix equal to the density matrix. For the TLQS, Slater [53, eq. (13)] [54, eq. (16)] derived the corresponding Fisher information metric. This is representable as,

$$ds^2_{Fisher_{Gauss}} = \frac{2(1+r^2)}{(1-r^2)^2}dr^2 + \frac{2}{1-r^2}dn^2. \tag{36}$$

The tangential component can be reproduced, following the basic formula (1), by choosing $f_{Gauss}(t) = \frac{t}{1+t}$. This is simply *one-half* of that — $f_{YL}(t) = 2f_{Gauss}(t) = \frac{2t}{1+t}$ — associated

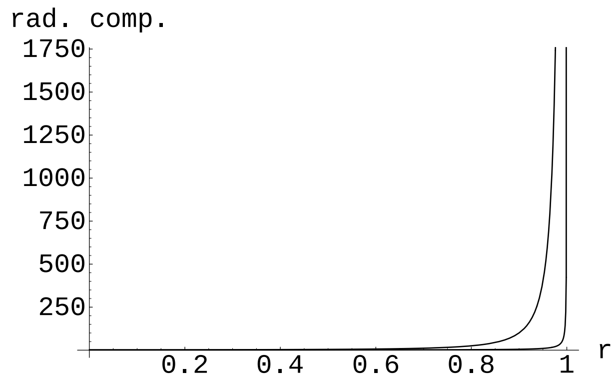

FIG. 8: Radial components of $ds^2_{monotone}$ and $ds^2_{Fisher_{Gauss}}$. The latter dominates the former.

with the *maximal* monotone (Yuen-Lax) metric [55]. Like that metric, the metric (36) yields a *non*-normalizable volume element (so one can not immediately apply — without some preliminary truncation — the comparative noninformativity/relative entropy test we have used above [25, 26]). Of course, the radial component of (36) is also not consistent with the requirement for a monotone metric. In fact, it rises much *more* steeply than $\frac{1}{1-r^2}$, in opposite behavior to that for $ds^2_{Fisher_{Hus}}$. In Fig. 8 we show this phenomenon.

VIII. DISCRETE WIGNER FUNCTION FOR A QUBIT

The discrete Wigner function (pseudoprobability) W, in the simplest case of a qubit, is defined on a 2×2 array, with four components $W_{i,j}, i, j = 1, 2$ [56, eqs. (14)-(17)]. The sum of W_{ij} in each "line" λ is the probability p_{ij} of projecting the state onto the basis vector $|\alpha_{ij} >$, where $i \in \{1, 2, 3\}$ indexes a set of three mutually unbiased bases (MUB) for a qubit and $j \in \{1, 2\}$ indexes the basis vector in each MUB. Choosing the MUB to be the eigenstates of the three Pauli operators, and using our cartesian coordinates, one can obtain three one-dimensional *marginal* (binomial) probability distributions over the $x-, y-$ and z-axes, of the form $(\frac{1+x}{2}, \frac{1-x}{2}), \cdots$ (cf. [57, 58]). Now, the corresponding Jeffreys' prior for the one-dimensional family of such binomial distribution is simply the *beta* distribution $p_\beta(x) = \frac{1}{\pi\sqrt{1-x^2}}$. (Let us note that the one-dimensional marginal distributions obtained for p_{Bures} are of another form, that is, $\frac{2\sqrt{1-x^2}}{\pi}$.)

Let us take the product of $p_\beta(x), p_\beta(y)$ and $p_\beta(z)$, which naturally forms a (prior) prob-

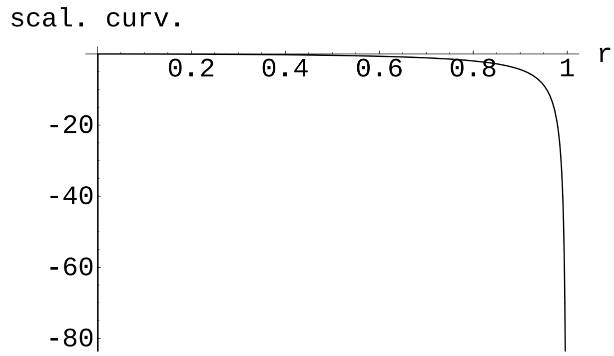

FIG. 9: Scalar curvature of the Fisher information metric for the family of Husimi distributions

ability distribution,

$$p_{product} = \frac{1}{\pi^3 \sqrt{(1-x^2)(1-y^2)(1-z^2)}}, \tag{37}$$

over the *hypercube* with vertices $(\pm 1, \pm 1, \pm 1)$ and renormalize/truncate it to a probability distribution over the Bloch sphere,

$$p_{Wigner} = \frac{1}{6.61455516101 \sqrt{(1-x^2)(1-y^2)(1-z^2)}}. \tag{38}$$

(Thus, the quantum-mechanically *inaccessible* region lying outside the Bloch ball, but within the hypercube is disregarded — assigned null measure — in the new normalization.)

Now, we found — strictly following the notation, formulas and line of argument above in sec. III — that $S_{KL}(p_{Wigner}, p_{Hus}) = 0.0149831$ and $S_{KL}(p_{Hus}, p_{Wigner}) = 0.0156225$, so these two distributions are rather close in nature. Of course, p_{Hus} is rotationally-symmetric over the Bloch sphere, while p_{Wigner} is not, so it seems to make little sense to try to compute some function $f_{Wigner}(t)$ to generate the tangential component. We found it problematical, using our usual (relative entropy) approach, to designate either p_{Hus} or p_{Wigner} as more or less noninformative. (The "Husimi function is a kind of...coarse-grained Wigner function" [48, p. 3].)

IX. SCALAR CURVATURE

In Fig. 9, we plot the *scalar curvature* of $ds^2_{Fisher_{Hus}}$. The formula for this scalar curvature

is

$$K^{n=2}_{Hus} = \frac{r\,\left(-6\,r + W\,\left(-3 + r^2\right)\right)\,\left(-4\,r^2\,\left(-3 + r^2\right) + 6\,W\,r\,\left(2 - 3\,r^2 + r^4\right) + W^2\,\left(3 - 8\,r^2 + 5\,r^4\right)\right)}{\left(W + 2\,r\right)^2\,\left(-1 + r^2\right)\,\left(-2\,r + W\,\left(-1 + r^2\right)\right)^2}, \tag{39}$$

where $W = \log\frac{1-r}{1+r}$. Also, expanding about $r = 0$,

$$K^{n=2}_{Hus} \approx \frac{-6\,r^2}{5} - \frac{138\,r^4}{125} - \frac{32094\,r^6}{30625} - \frac{154474\,r^8}{153125} - \frac{57710054\,r^{10}}{58953125}. \tag{40}$$

The *nonpositive* monotonically-*decreasing* scalar curvature (Fig. 9) has its *maximum* at $r = 0$, corresponding to the fully mixed state, indicative of a *flat* metric there (cf. (7)) (and is $-\infty$ at the pure states, $r = 1$). For the minimal monotone (Bures) metric, the *nonnegative* scalar curvature is *constant*, that is $K^{n=2}_{min} = 6$, over the Bloch ball, and for the (n^2-1)-dimensional convex set of $n \times n$ density matrices, $n > 2$, achieves its *minimum* of $K^n_{min} = \frac{(5n^2-4)(n^2-1)}{8}$ at the fully mixed state ($\rho = \frac{1}{n}I$) [59]. (In [59], the metric used is one-quarter of that corresponding to (1), used here, so the results we compute here differ from those there by such a factor. For the *maximal* monotone metric, $K^{n=2}_{max} = \frac{8(r^2-6)}{1-r^2}$, which is monotonically-*decreasing* as r increases, as is $K^{n=2}_{Hus}$.)

For the two-level quantum sytems, Andai [60] has constructed a family of monotone metrics with *non*-monotone scalar curvature, and given a condition for a monotone metric to have a local minimum at the maximally mixed state.

A. Metrics of constant scalar curvature

The metric $ds^2_{Fisher_{q=2}}$ has *constant* scalar curvature, $K^{n=2}_{q=2} = \frac{3}{2}$, while, as previously noted, $K^{n=2}_{min} = 6$. Let us note that $K^n_{WY} = \frac{1}{4}(n^2-1)(n^2-2)$, which is also $\frac{3}{2}$ for $n = 2$. Here, WY denotes the Wigner-Yanase metric — the only pull-back metric among the quantum monotone metrics — and $f_{WY}(t) = \frac{1}{4}(\sqrt{t}+1)^2$, which is the only *self-dual* operator monotone function [61]. "It is not known at the moment if there are other monotone metrics of constant sectional and scalar curvature" [61, p. 3760]. It is a theorem that the "set of two-dimensional normalized density matrices equipped with the Bures metric is isometric to one closed-half of the three-sphere with radius $\frac{1}{2}$" [62]. The WY-metric "looks locally like a sphere of radius 2 of dimension (n^2-1)" [61, p. 3759]. If we transform to spherical coordinates on the 3-sphere, then, the metric tensor for ds^2_{min} is diagonal in character, while the two other (constant scalar curvature) metrics are not (cf. [63]).

The three metrics ds^2_{min}, ds^2_{WY} and $ds^2_{Fisher_{q=2}}$ are *Einstein.* If we *scale* these metrics so that they are all of *unit* volume [64], then $K^{n=2}_{min/scaled} = 6\pi^2 \approx 59.2176, K^{n=2}_{WY/scaled} = 6\pi(\pi-2) \approx 21.5185$ and $K^{n=2}_{q=2/scaled} = 4\pi^2 - 6\sqrt{3}\pi \approx 6.83003$. The constant scalar curvatures of (unit-volume) *Yamabe* metrics are bounded above, and their least upper bound is a real number equal to $n(n-1)V_n^{2/n}$, where V_n is the volume of the standard metric on S^n, and in our (Bloch sphere) case, $n = 3$, so the bound is $242^{\frac{1}{3}}\pi^{\frac{4}{3}} \approx 139.13$ [64].

X. DISCUSSION

Luo [24] (cf. [44, sec. 2.4] [19, 20, 65]) has calculated the Fisher information matrix of the Husimi distribution in the Fock-Bargmann representation of the quantum harmonic oscillator with one degree of freedom. He found that the Fisher information of the position and that of the momentum move in opposite directions, and that a weighted trace of the Fisher information matrix is a constant independent of the wave function, and thus has an upper bound. (Luo did not consider the possibility of generating prior probability distributions by normalizing the volume element of the Fisher information metric.)

Gnutzmann and Życzkowski noted that one "is tempted to think of the Husimi function as a probability density on the phase space. However, the rules for calculating expectation values of some observable using the Husimi function are non-classical" [47, sec. 2.1] (cf. [66, p. 548]). Gardiner and Zoller remarked that the "main problem of the Q-function is that not all positive normalizable Q-functions correspond to positive normalizable density operators" [51, p. 109].

Further, the comparison of distances between Husimi distributions for arbitrary quantum states based on the Fisher information metric with those employing the Monge distance [3], might be investigated. For the TLQS studied here, the Monge distance is, in fact, "consistent with the geometry of the Bloch ball induced by the Hilbert-Schmidt or the trace distance" [3, p. 6716]. (The trace distance is monotone, but *not* Riemannian, while the Hilbert-Schmidt distance, contrastingly, is Riemannian, but *not* monotone [67, p. 10083] [37].) For n-dimensional quantum systems ($n > 2$), unlike the trace, Hilbert-Schmidt or Bures distance, the Monge distance of ρ to the fully mixed state — which provides information concerning the *localization* of ρ in the classical phase space — is *not* the same for all pure states [3]. The only monotone metrics for which explicit distance formulas are so-far available are the

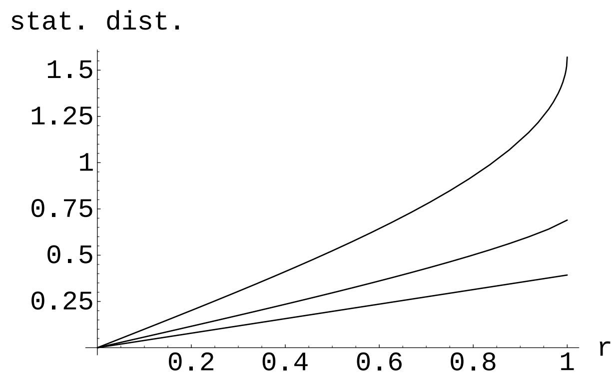

FIG. 10: Statistical distance as a function of distance from the origin of the Bloch ball — corresponding to the fully mixed state — for any monotone metric, for $ds^2_{Fisher_{Hus}}$, and for the Monge (or equivalently, for $n = 2$, Hilbert-Schmidt) metric. The monotone-metric curve dominates that for $ds^2_{Fisher_{Hus}}$, which dominates the linear curve for the Monge metric.

Bures (minimal monotone) and Wigner-Yanase ones [61].

In Fig. 10 we show how the distance from the fully mixed state ($r = 0$) increases as r increases, for any monotone metric and for $ds^2_{Fisher_{Hus}}$, and (linearly) for the Monge (or Hilbert-Schmidt) metric. The first-mentioned distance — taking the functional form $\arcsin r$ (equalling $\frac{\pi}{2}$ for $r = 1$) — dominates the second-mentioned distance (equalling $\frac{\pi}{4.5551532167057}$ for $r = 1$), which in turns dominates the third [3, eq. (4.10)], which takes the value $\frac{\pi}{8}$ for $r = 1$.

Let us bring to the attention of the reader, a recent preprint, which introduces a concept of escort *density operators* and a related one of *generalized* Fisher information [68] (cf. [46, 69]).

We have been consistently able above to find (apparently operator) monotone functions to generate the tangential components of (classical) Fisher information metrics for (rotationally-symmetric) probability distributions over the TLQS. We suspect the existence of some (yet not formally demonstrated) theorem to this effect. Also, it would be of interest to formally test the various monotone functions presented above for the property (requisite for a *quantum* monotone metric [15, 21] of *operator* monotonicity.

We have "hybridized" $ds^2_{Fisher_{Hus}}$ above to a (quantum) monotone metric $ds^2_{HYB_{Hus}}$ by replacing its radial component by that required ($\frac{1}{1-r^2}$), while retaining its tangential component (formed from $f_{Hus}(t)$). But it appears that we could also convert it by appropriately scaling (a *conformal* transformation) the entire metric (tangential *and* radial components)

by some suitable function. If we do so, we find that — by explicit construction — the new metric ($ds^2_{conformal_{Hus}}$) has the required radial component, while the tangential component is generated by a function

$$f_{conformal_{Hus}}(t) = f_{Hus}(t) - t - 1, \tag{41}$$

which also appears to be operator monotone. (We note that $f_{conformal_{Hus}}(1) = 1$ and $\lim_{t\to 0} f_{conformal_{Hus}}(t) = 0$.) But now, we have the large relative entropies $S_{KL}(p_{GKS}, p_{conformal_{Hus}}) = 50.4636$ and $S_{KL}(p_{conformal_{Hus}}, p_{GKS}) = 54.2601$. At $r = 0$, $ds^2_{conformal_{Hus}}$ is not flat, as is $ds^2_{Fisher_{Hus}}$, but has a (limiting) scalar curvature of $-\frac{24}{5}$.

A. Further questions

Motivated by the analyses above, we would like to pose the question of whether there exists a family of trivariate *probability* distributions parametrized by the points of the Bloch ball, for which the associated (classically *unique* [up to a constant multiple]) Fisher information metric *fully* — both in terms of tangential *and* radial components – has the requisite form (1) for a monotone metric. Also, the *volume elements* (and hence associated prior probabilities) of the monotone metrics are expressible as the *product* of Haar measure and measures over the eigenvalues [16]. To what extent, if any, does this hold true for prior probabilities *not* arising from monotone metrics? Are there any non-monotone metrics which give rise to prior probabilities *more* noninformative than (at the very least) the minimal monotone (Bures) one? What are suitable counterparts to formula (1) for n-level quantum systems ($n > 2$)? Are there any monotone metrics which are flat at the fully mixed state, as is $ds^2_{Fisher_{Hus}}$ (7)?

XI. SUMMARY

In a *classical* context, for the family of Husimi *probability* distributions over the three-dimensional Bloch ball of two-level quantum systems (TLQS), we derived the (flat-at-the-fully-mixed-state) Fisher information metric ($ds^2_{Fisher_{Hus}}$, given by (2)). Its tangential — but *not* its radial (r) — component conformed to that of one of the (uncountably) *infinite* class of (quantum) monotone metrics. The *prior* probability distribution (p_{Hus}) formed by nor-

malizing the volume element of $ds^2_{Fisher_{Hus}}$ was found (sec. III A) to be considerably *less* non-informative than the priors formed from *any* of the (quantum) monotone metrics, even that (p_{Bures}) based on the (relatively informative) *minimal* monotone (Bures) metric. However, if we replaced the radial component of $ds^2_{Fisher_{Hus}}$ by that required ($\frac{1}{1-r^2}$) of *all* (quantum) monotone metrics, the resultant "hybridized-Husimi" prior probability ($p_{HYB_{Hus}}$) became very close (in the sense of relative entropy $\approx .006$ "nats") to p_{Bures}, and thus comparably informative in nature, but even nearer ($\approx .0004$) to another quantum-monotone-metric-based ("Grosse-Krattenthaler-Slater" or "quasi-Bures") probability distribution (p_{GKS}) that has been conjectured to yield the asymptotic minimax/maximin redundancy for universal *quantum* coding. The analogous (Bayesian) role in universal (classical) coding — by a well-known result of Clarke and Barron [17, 18] — is played by Jeffreys' prior (cf. [19, 20]). This takes the specific (original, *non*-hybridized) form p_{Hus} for the family (manifold) of trivariate Husimi qubit probability distributions under study. We also studied the Fisher information metric for the *escort*-Husimi (sec. V), positive-P (sec. VI) and certain Gaussian probability distributions (sec VII), as well as, in some sense, the discrete Wigner pseudoprobability (sec. VIII). Additionally, we applied the Clarke comparative noninformativity test [25, 26] to quantum priors (sec. III). Evidence that this test is consistent with the recently-stated criterion of "biasedness to pure states" of Srednicki [27] has been presented [28].

Acknowledgments

I wish to express gratitude to the Kavli Institute for Theoretical Physics (KITP) for computational support in this research and to C. Krattenthaler for deriving the expressions (30) and (33) and to M. Trott for assistance with certain MATHEMATICA computations.

---

[1] R. Mosseri and R. Dandoloff, J. Phys. A. **34**, 10243 (2001).

[2] K. Husimi, Proc. Phys. Soc. Japan **22**, 264 (1940).

[3] K. Życzkowski and W. Słomczyński, J. Phys. A **34**, 6689 (2001).

[4] S. Guiasu, Phys. Rev. A **36**, 1971 (1987).

[5] R. D. Levine, J. Statist. Phys. **52**, 1203 (1988).

[6] A. Bach, Phys. Lett. A **73**, 287 (1979).

[7] A. Bach, J. Math. Phys. **21**, 789 (1980).

[8] A. Bach, J. Phys. A **14**, 125 (1981).

[9] G. W. Gibbons, J. Geom. Phys. **8**, 147 (1992).

[10] A. Sugita and H. Aiba, Phys. Rev. E **65**, 036205 (2002).

[11] R. Penrose, *The Road to Reality* (A. A. Knopf, New York, 2005).

[12] N. N. Chentsov, *Statistical Decision Rules and Optimal Inference* (Amer. Mat. Soc., Providence, 1982).

[13] L. L. Campbell, Proc. Amer. Math. Soc. **98**, 135 (1986).

[14] B. R. Frieden, *Science from Fisher Information: A Unification* (Cambridge Univ, Cambridge, 2004).

[15] D. Petz and C. Sudár, J. Math. Phys. **37**, 2662 (1996).

[16] M. J. W. Hall, Phys.Lett.A **242**, 123 (1998).

[17] B. S. Clarke and A. R. Barron, IEEE Trans. Info. Theory **36**, 453 (1990).

[18] B. S. Clarke and A. R. Barron, J. Statist. Plann. Inference **41**, 37 (1994).

[19] P. B. Slater, J. Phys. A **29**, L601 (1996).

[20] L. C. Kwek, C. H. Oh, and X.-B. Wang, J. Phys. A **32**, 6613 (1999).

[21] D. Petz, Linear Alg. Applics. **244**, 81 (1996).

[22] P. B. Slater, Lett. Math. Phys. **52**, 343 (2000).

[23] D. Petz, J. Phys. A **35**, 929 (2002).

[24] S. Luo, J. Statist. Phys. **102**, 1417 (2001).

[25] P. B. Slater, Phys. Lett. A **247**, 1 (1998).

[26] B. Clarke, J. Amer. Statist. Assoc. **91**, 173 (1996).

[27] M. Srednicki, Phys. Rev. A **71**, 052107 (2005).

[28] P. B. Slater, quant-ph/0507203 (to appear in J. Phys. A).

[29] P. B. Slater, J. Phys. A **29**, L271 (1996).

[30] P. B. Slater, J.Math. Phys. **37**, 2682 (1996).

[31] F. Mintert and K. Życzkowski, Phys. Rev. A **69**, 022317 (2004).

[32] A. Lesniewski and M. B. Ruskai, J. Math. Phys. **40**, 5702 (1999).

[33] A. Jenčová, Intl. J. Theor. Phys. **43**, 1635 (2004).

[34] S. Luo, Proc. Amer. Math. Soc. **132**, 885 (2003).

[35] P. B. Slater, J. Phys. A **34**, 7029 (2001).

[36] T. Decker, D. Janzing, and T. Beth, Intl. J. Quant. Inform. **2**, 353 (2004).

[37] M. Ozawa, Phys. Lett. A **268**, 158 (2001).

[38] C. Krattenthaler and P. B. Slater, IEEE Trans. Info. Theory **46**, 801 (2000).

[39] M. Hayashi and K. Matsumoto, Phys. Rev. A **66**, 022311 (2002).

[40] R. Jozsa and S. Presnell, Proc. Royal Soc. Lond. A **459**, 3061 (2003).

[41] R. Jozsa, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **81**, 1714 (1998).

[42] H. Grosse, C. Krattenthaler, and P. B. Slater, *Asymptotic redundancies for universal quantum coding. part 2 (in preparation).*

[43] C. M. Caves, C. A. Fuchs, and R. Schack, J. Math. Phys. **43**, 4537 (2002).

[44] F. Pennini and A. Plastino, Phys. Lett. A **326**, 20 (2004).

[45] S. I. Amari, Ann. Statist. **10**, 357 (1982).

[46] J. Naudts, J. Inequal. Pure Appl. Math. **5**, 1 (2004).

[47] S. Gnutzmann and K. Życzkowski, J. Phys. A **34**, 10123 (2001).

[48] A. Sugita, J. Phys. A **36**, 9081 (2003).

[49] C. Beck, Phys. Lett. A **287**, 240 (2001).

[50] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Phys. Rev. A **43**, 1153 (1991).

[51] C. W. Gardiner and P. Zoller, *Quantum Noise* (Springer, Berlin, 2004).

[52] H. W. Lee, Phys. Rep. **259**, 147 (1995).

[53] P. B. Slater, Phys. Lett. A **206**, 66 (1995).

[54] P. B. Slater, Phys. A **223**, 167 (1996).

[55] H. P. Yuen and M. Lax, IEEE Trans. Info. Theory **19**, 740 (1973).

[56] E. F. Galvão, Phys. Rev. A **71**, 042302 (2005).

[57] R. Asplund and G. Björk, Phys. Rev. A **64**, 012106 (2001).

[58] M. O. T. Cunha, V. I. Man'ko, and M. O. Scully, Found. Phys. Lett. p. 103 (2001).

[59] J. Dittmann, J. Geom. Phys. **31**, 16 (1999).

[60] A. Andai, J. Math. Phys. **44**, 3675 (2003).

[61] P. Gibilisco and T. Isola, J. Math. Phys. **44**, 3752 (2003).

[62] M. Hübner, Phys. Lett. A **163**, 239 (1992).

[63] L. Molnár and W. Timmermann, J. Phys. A **36**, 267 (2003).

[64] M. J. Gursky and C. LeBrun, GAFA Geom. Funct. Anal. **8**, 965 (1998).

[65] J. Twamley, J. Phys. A **29**, 3723 (1996).

[66] G. Esposito, G. Marmo, and G. Sudarshan, *From Classical to Quantum Mechanics* (Cambridge Univ. Press, Cambridge, 2004).

[67] H.-J. Sommers and K. Życzkowski, J. Phys. A **36**, 10083 (2003).

[68] J. Naudts, Open Syst. Inform. Dyn. **12**, 13 (2005).

[69] E. Lutwak, D. Yang, and G. Y. Zhang, IEEE Trans. Info. Theory **51**, 473 (2005).